RECEIVED
CENTRAL FAX CENTER

AUG 2 4 2006

AMENDMENT UNDER 37 C.F.R. § 1.111
U.S. Application No. 10/729,915

AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions and listings of claims in the application:

LISTING OF CLAIMS:

1. (*Currently Amended*) A method for organizing a plurality of data files using meta-data ~~meta-data, having~~ wherein at least one ~~meta-data~~ meta-data element ~~at least~~ is associated with each data file, the method comprising:

extracting, for at least some ~~of the~~ selected data files, at least one meta-data element associated with ~~that data file~~ a respective selected data file;

organizing the extracted meta-data elements in a desired order ~~based on values for the extracted meta-data elements~~;

inputting at least one ~~parameter~~ value of a clustering sensitivity parameter for determining clustering based on pair-wise comparisons between values of the extracted meta-data elements; and

dividing ~~at least some of~~ the selected data files into groups based on the extracted meta-data elements associated with the selected data files and the ~~input~~ at least one value of the clustering sensitivity parameter value.

2. (*Currently Amended*) The method of claim 1, wherein dividing the ~~at least some~~ selected data files into groups comprises determining ~~determining, for each of at least one of the at least one parameter value,~~ a similarity value for at least ~~two~~ one pair of the selected ~~plurality of~~

RECEIVED
CENTRAL FAX CENTER

AUG 2 4 2006

data files for the value of the clustering sensitivity parameter, wherein the similarity value of the one pair of the selected data files depends on the value of the clustering sensitivity parameter and the values ~~using at least some~~ of the extracted meta-data elements of the at least one pair of the selected data files ~~and that parameter value~~.

3. (*Currently Amended*) The method of claim 2, wherein determining the ~~at least one~~ similarity value of the one pair comprises ~~determining the at least one~~ evaluating the similarity value as:

$$S_K(i,j) = \exp\left(-\frac{|t_i - t_j|}{K}\right),$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value; ~~and~~

~~t_i and t_j are actual values of one extracted meta-data element for the i^{th} and j^{th} data files~~

t_i is the value of a selected extracted meta-data element associated with the i^{th} data file; and

t_j is the value of a selected extracted meta-data element associated with the j^{th} data file.

RECEIVED
CENTRAL FAX CENTER

AUG 2 4 2006

4. (*Currently Amended*) The method of claim 2, wherein determining the ~~at least one~~ similarity value of the one pair comprises determining the ~~at least one~~ similarity value of the one pair as :

$$S_K(i,j) = \exp\left(\frac{1}{K}\left(-\frac{< v_i, v_j >}{|v_i||v_j|} - 1\right)\right) .$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value; and

v_i and v_j are actual vector values determined from the i^{th} and the j^{th} data files.

5. (*Currently Amended*) The method of claim 2, further comprising determining, for each of the ~~at least some~~ selected data files, at least one novelty value for that each data file based on ~~the~~ at least one similarity value ~~for~~ comprising that each data file and ~~for a number of~~ a nearby data files.

6. (*Original*) The method of claim 5, wherein determining at least one novelty value comprises determining at least one novelty value as:

$$v_K(s) = \sum_{l,n=-5}^{5} S_K(s+1, s+n)\, g(1,n) .$$

where:

$v_K(s)$ is the novelty value; and

g is a Gaussian tapered 11 x 11 checkerboard kernel.

7. (*Currently Amended*) The method of claim 5, further comprising determining at least one boundary location between the ~~ones of the plurality of~~ selected data files based on the at least one novelty values determined for the each ~~at least some~~ of the selected data files.

8. (*Currently Amended*) The method of claim 7, further comprising determining, for at least ~~some~~ the one determined boundary locations, a confidence value for that boundary location.

9. (*Currently Amended*) The method of claim 8, wherein determining a confidence value for a boundary location comprises determining the confidence value as:

$$C(B_K) = \sum_{l=1}^{|B_K|-1} \frac{1}{(b_{l+1} - b_1)^2} \sum_{i,j=b1}^{b_{l+1}} S_K(i,j) - \sum_{l=1}^{|B_K|-2} \frac{1}{(b_{l+1} - b_1)(b_{l+2} - b_{l+1})} \sum_{i=b_1}^{b_{l+1}} \sum_{j=b_{l+1}}^{b_{l+2}} S_K(i,j) \ .$$

where:

$C(B_K)$ is the confidence value for the B_K^{th} boundary;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

b is the index value of detected boundary at a particular value for the input clustering sensitivity parameter K level.

10. (*Currently Amended*) The method of claim 8, further comprising determining, for at least ~~some of~~ the one determined boundary locations, ~~at least one of the~~ at least one clustering sensitivity parameter value that maximizes the confidence value.

11. (*Currently Amended*) A method for organizing a plurality of data files using meta-data ~~having~~, wherein at least one meta-data element ~~that~~ is at least associated with a corresponding one of the plurality of data files, the method comprising:

processing at least one set of meta-data elements, where each meta-data element of the set corresponds to a data file;

obtaining a desired value of a clustering sensitivity parameter for analyzing the meta-data; and

determining a structure within the processed set of the meta-data elements ~~using an obtained parameter value, wherein the structure is determined~~ by comparing, for at least a subset of the plurality of data files, elements of at least a subset of the processed meta-data elements ~~using the parameter value~~ to each other using the obtained value of the clustering sensitivity parameter.

12. (*Currently Amended*) The method of claim 11, further comprising clustering the data files into groups using the determined structure of the meta-data elements.

13. (*Original*) The method of claim 12, further comprising determining boundaries from the determined clusters of data files, wherein the boundaries are located between the determined clusters of data files.

14. (*Currently Amended*) The method of claim 13, further comprising:

determining a similarity value by comparing at least some of the meta-data elements ~~in one~~ associated with data files in a first selected cluster of data files to at least some other ~~ones of the meta-data~~ meta-data elements associated with data files of in ~~that element~~ the first selected cluster of data files; and

determining a dissimilarity value by comparing at least some of the meta-data elements associated with data files in ~~one~~ a second selected cluster of data files to at least some of the meta-data elements associated with data files in ~~another~~ a third selected cluster of data files, wherein the second selected cluster of data files and the third selected cluster of data files are different clusters of data files.

15. (*Original*) The method of claim 14, further comprising:

determining a value corresponding to a desired grouping of the clusters of data files based on the differences of the similarity values and the dissimilarity values.

16. (*Currently Amended*) A storage medium storing a set of program instructions executable on a data processing device and usable to organize a plurality of data files by using meta-data ~~meta-data having~~ wherein at least one ~~meta-data~~ meta-data element ~~at least~~ is associated with each data file, the program comprising:

instructions for extracting, for at least some of the data files, at least one meta-data element associated with each of the some data files ~~that data file~~;

instructions for organizing the extracted meta-data elements in a desired order ~~based on values for the extracted meta-data elements~~;

instructions for inputting ~~a~~ at least one clustering sensitivity parameter value; and

instructions for dividing the ~~at least~~ some ~~of the~~ data files into groups based on the extracted meta-data elements associated with the some data files and the at least one ~~input~~ clustering sensitivity parameter value.

17. (*Currently Amended*) The storage medium of claim 16, the instructions for dividing the at least some ~~of the~~ data files into groups further comprising instructions for determining, ~~for each of at least of the at least one parameter value,~~ a similarity value for at least ~~two~~ one pair of the some ~~plurality of~~ data files for the value of the clustering sensitivity parameter, wherein the similarity value of the one pair of the some data files depends on the value of the clustering sensitivity parameter and the values ~~using at least some~~ of the extracted meta-data elements of the at least one pair of the some data files ~~and that parameter value~~.

18. (*Currently Amended*) The storage medium of claim 17, further comprising instructions for determining, for each of the at least some data files, at least one novelty value for that each data file based on ~~the~~ at least one similarity value ~~for~~ comprising that each data file and ~~for a number of~~ a nearby data files.

19. (*Currently Amended*) The storage medium of claim 17, wherein instructions for determining the ~~at least one~~ similarity value of the one pair comprises instructions for ~~determining~~ evaluating the ~~at least one~~ similarity value as:

$$S_K(i,j) = \exp\left(-\frac{|t_i - t_j|}{K}\right),$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value;

t_i is the value of a selected extracted meta-data element associated with the i^{th} data file; and

t_j is the value of a selected extracted meta-data element associated with the j^{th} data file.

~~t_i and t_j are actual values of the at least one meta-data element of the at least one extracted meta-data element for the i^{th} and j^{th} data files.~~

20. (*Currently Amended*) The storage medium of claim 17, wherein instructions for determining the ~~at least one~~ similarity value of the one pair comprises determining the ~~at least one~~ similarity value of the one pair as:

$$S_K(i,j) = \exp\left(\frac{1}{K}\left(-\frac{< v_i - v_j >}{|v_i||v_j|} - 1\right)\right).$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value; and

v_i and v_j ~~that~~ are actual vector values determined from the i^{th} and the j^{th} data files.

21. (*Currently Amended*) The storage medium of claim 18, further comprising instructions for determining at least one boundary location between ~~ones of the plurality of~~ the some data files based on the at least one novelty values determined for the each of the at least some ~~of the~~ data files.

22. (*Original*) The storage medium of claim 18, wherein instructions for determining at least one novelty value comprises instructions for determining the at least one novelty value as:

$$v_K(s) = \sum_{l,m=-5}^{5} S_K(s+1, s+n)g(1,n) \ .$$

where:

$v_K(s)$ is the novelty value; and

g is the Gaussian tapered 11 x 11 checkerboard kernel.

23. (*Currently Amended*) The storage medium of claim 21, further comprising instructions for determining, for at least ~~some of~~ the one determined boundary locations, a confidence value for that boundary location.

24. (*Original*) The storage medium of claim 23, wherein instructions for determining at least one confidence value comprises instructions for determining each of such confidence value as:

$$C(B_K)=\sum_{l=1}^{|B_K|-1}\frac{1}{(b_{l+1}-b_1)^2}\sum_{i,j=b1}^{b_{l+1}}S_K(i,j)-\sum_{l=1}^{|B_K|-2}\frac{1}{(b_{l+1}-b_1)(b_{l+2}-b_{l+1})}\sum_{i=b_1}^{b_{l+1}}\sum_{j=b_{l+1}}^{b_{l+2}}S_K(i,j)\ .$$

where:

$C(B_K)$ is the confidence value for the B_K^{th} boundary;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

b is the detected boundary at a level.

25. (*Currently Amended*) The storage medium of claim 23, further comprising instructions for determining, for at least ~~some of the~~ the one determined boundary locations, ~~at least one of the~~ at least one clustering sensitivity parameter value that maximizes the confidence value.

26. (*Currently Amended*) A data file organizing system usable to organize a plurality of data files using meta data having at least one meta data element that is at least associated with a corresponding one of the data files, comprising:

a meta-data extracting circuit, routine, or application that extracts, for at least some of the data files, at least one meta-data element associated with that data file;

a meta-data organizing circuit, routine or application that organizes the extracted meta-data elements in a desired order based on values for the extracted meta-data elements;

a similarity value determining circuit, routine or application that determines, for at least one of the at least one clustering sensitivity parameter value, a similarity value for at least two of the plurality of data files using at least some of the extracted meta-data elements and that clustering sensitivity parameter value

a novelty value determining circuit, routine or application that determines at least one novelty value for that data file based on the at least one similarity value for that data file and for a number of nearby data files;

a data dividing determining circuit, routine or application that divides at least some of the data files into groups based on the extracted meta-data elements and the input clustering sensitivity parameter value by determining at least one boundary location between ones of the plurality of data files based on the at least one novelty value determined for at least some of the data files; and

a confidence value determining circuit, routine or application that determines, for at least some of the determined boundary locations, a confidence value for that boundary location, wherein the data dividing circuit, routine, or application further determines, for at least some of the determined boundary locations, the at least one clustering sensitivity parameter value that maximizes the confidence value.

27. (*New*) The method of claim 2, wherein determining the similarity value of the one pair comprises a term depending on a scalar magnitude of a difference between t_i and t_j relative to K, where K is the clustering sensitivity parameter value, i references a first data file of the one pair, j references a second file of the one pair, t_i is the value of a selected extracted meta-

data element associated with the first data file, and t_j is the value of a selected extracted meta-data element associated with the second data file.

28. (*New*) The method of claim 27 wherein the depending comprises an exponentially decreasing function of the scalar magnitude of the difference between t_i and t_j relative to K.

29. (*New*) The method of claim 2, wherein determining the similarity value of the at least one pair of the selected data files comprises a term depending on an inner product of v_i and v_j relative to K, where K is the clustering sensitivity parameter value, v_i is an actual vector value determined from the i^{th} data file, and v_j is an actual vector value determined from the j^{th} data file.